Amendments to the Specification

Applicants submit a substitute specification herewith. Applicants have included both a marked up version and a clean copy of the specification. The substitute specification contains no new matter.

TITLE OF THE INVENTION

AGGRECANASE MOLECULES

This application is a continuation-in-part of USSN 09/978,979 filed October 16, 2001.

The present invention relates to the discovery of nucleotide sequences encoding novel aggrecanase molecules, the aggrecanase proteins and processes for producing them. The invention further relates to the development of inhibitors of, as well as antibodies to, the aggrecanase enzymes. These inhibitors and antibodies may be useful for the treatment of various aggrecanase-associated conditions including osteoarthritis.

BACKGROUND OF THE INVENTION

Aggrecan is a major extracellular component of articular cartilage. It is a proteoglycan responsible for providing cartilage with its mechanical properties of compressibility and elasticity. The loss of aggrecan has been implicated in the degradation of articular cartilage in arthritic diseases. Osteoarthritis is a debilitating disease, which affects at least 30 million Americans [MacLean et al. J Rheumatol 25: 2213-8 (1998)]. Osteoarthritis can severely reduce the quality of life due to degradation of articular cartilage and the resulting chronic pain. An early and important characteristic of the osteoarthritic process is loss of aggrecan from the extracellular matrix [Brandt, KD. and Mankin, HJ. Pathogenesis of Osteoarthritis, Textbook of Rheumatology, WB Saunders Company, Philadelphia, PA: 1355-1373 (1993)]. The large, sugar-containing portion of aggrecan is thereby lost from the extracellular matrix, resulting in deficiencies in the biomechanical characteristics of the cartilage.

A proteolytic activity termed "aggrecanase" is thought to be responsible for the cleavage of aggrecan thereby having a role in cartilage degradation associated with osteoarthritis and inflammatory joint disease. Work has been conducted to identify the enzyme responsible for the degradation of aggrecan in human osteoarthritic cartilage. Two enzymatic cleavage sites have been identified within the interglobular domain of aggrecan. One (Asn^{341} -Phe^{342}) is observed to be cleaved by several known metalloproteases [Flannery, CR et al. J Biol Chem 267: 1008-14

(1992); Fosang, AJ et al. Biochemical J. 304: 347-351. (1994)]. The aggrecan fragment found in human synovial fluid, and generated by IL-1 induced cartilage aggrecan cleavage, is at the Glu^{373}-Ala^{374} bond [Sandy, JD et al. J Clin Invest 69: 1512-1516 (1992); Lohmander, LS et al. Arthritis Rheum 36: 1214-1222 (1993); Sandy, JD et al. J Biol Chem. 266: 8683-8685 (1991)], indicating that none of the known enzymes are responsible for aggrecan cleavage *in vivo*.

Recently, identification of two enzymes, aggrecanase-1 (ADAMTS4) and aggrecanase-2 (ADAMTS11), within A Disintegrin-like And Metalloprotease with ThromboSpondin type 1 motif (ADAM-TS) family have been identified which are synthesized by IL-1 stimulated cartilage and cleave aggrecan at the appropriate site [Tortorella, MD et al. Science 284: 1664-6 (1999); Abbaszade, I et al. J Biol Chem 274: 23443-23450 (1999)]. It is possible that these enzymes could be synthesized by osteoarthritic human articular cartilage. It is also contemplated that there are other, related enzymes in the ADAM-TS family, which are capable of cleaving aggrecan at the Glu^{373}-Ala^{374} bond and could contribute to aggrecan cleavage in osteoarthritis.

SUMMARY OF THE INVENTION

The present invention is directed to the identification of aggrecanase protein molecules capable of cleaving aggrecanase, the nucleotide sequences that encode the aggrecanase enzymes, and processes for the production of aggrecanase. These enzymes are contemplated to be characterized as having proteolytic aggrecanase activity. The invention further includes compositions comprising these enzymes as well as antibodies to these enzymes. In addition, the invention includes methods for developing inhibitors of aggrecanase, which block the enzyme's proteolytic activity. These inhibitors and antibodies may be used in various assays and therapies for treatment of conditions characterized by the degradation of articular cartilage.

The invention includes methods for obtaining the full-length aggrecanase molecule, the DNA sequence obtained by this method, and the protein encoded thereby. The method for isolation of the full-length sequence involves utilizing the aggrecanase sequence to design probes for screening using standard procedures known to those skilled in the art. One embodiment of the invention includes the full-length nucleotide sequence of an aggrecanase of the invention. This sequence is set forth in SEQ ID NO: 7 from nucleotide #1 through nucleotide #4284. This sequence encodes the amino acid sequence set forth in SEQ ID NO: 8 from amino acid #1

through amino acid #1427. The invention further includes fragments of SEQ ID NO: 8 encoding molecules, which exhibit aggrecanase activity.

It is expected that other species have DNA sequences homologous to human aggrecanase enzyme. The invention, therefore, includes methods for obtaining the DNA sequences encoding other aggrecanase molecules, the DNA sequences obtained by those methods, and the proteins encoded by those DNA sequences. This method entails utilizing the nucleotide sequence of the invention or portions thereof to design probes to screen libraries for the corresponding gene from other species or coding sequences or fragments thereof from using standard techniques. Thus, the present invention may include DNA sequences from other species, which are homologous to the human aggrecanase protein and can be obtained using the human sequence. The present invention may also include functional fragments of the aggrecanase protein, and DNA sequences encoding such functional fragments, as well as functional fragments of other related proteins. The ability of such a fragment to function is determinable by assay of the protein in the biological assays described for the assay of the aggrecanase protein.

In one embodiment, the aggrecanase protein of the invention may be produced by culturing a cell transformed with the DNA sequence of SEQ ID NO: 7 from nucleotides #1 to #4284 and recovering and purifying the aggrecanase protein comprising an amino acid sequence of SEQ ID NO: 8 from the culture medium. The purified expressed protein is substantially free from other proteinaceous materials with which it is co-produced, as well as from other contaminants. For production in mammalian cells, the DNA sequence further comprises a DNA sequence encoding a suitable propeptide 5' to and linked in frame to the nucleotide sequence encoding the aggrecanase enzyme.

The recovered purified protein is contemplated to exhibit proteolytic aggrecanase activity cleaving aggrecan. Thus, the proteins of the invention may be further characterized by the ability to demonstrate aggrecan proteolytic activity in an assay, which determines the presence of an aggrecan-degrading molecule. These assays or the development thereof is within the knowledge of one skilled in the art. Such assays may involve contacting an aggrecan substrate with the aggrecanase molecule and monitoring the production of aggrecan fragments [see for example, Hughes et al., Biochem J 305: 799-804 (1995); Mercuri et al., J. Bio Chem. 274: 32387-32395 (1999)].

In another embodiment, the invention includes methods for developing inhibitors of aggrecanase and the inhibitors produced thereby. These inhibitors prevent cleavage of aggrecan. The method may entail the determination of binding sites based on the three-dimensional structure of aggrecanase and aggrecan and developing a molecule reactive with the binding site. Candidate molecules are assayed for inhibitory activity. Additional standard methods for developing inhibitors of the aggrecanase molecule are known to those skilled in the art. Assays for the inhibitors involve contacting a mixture of aggrecan and the inhibitor with an aggrecanase molecule followed by measurement of the aggrecanase inhibition, for instance by detection and measurement of aggrecan fragments produced by cleavage at an aggrecanase susceptible site.

Another aspect of the invention provides pharmaceutical compositions containing a therapeutically effective amount of aggrecanase inhibitors in a pharmaceutically acceptable vehicle.

Aggrecanase-mediated degradation of aggrecan in cartilage has been implicated in osteoarthritis and other inflammatory diseases. Therefore, these compositions of the invention may be used in the treatment of diseases characterized by the degradation of aggrecan and/or an upregulation of aggrecanase. The compositions may be used in the treatment of these conditions or in the prevention thereof.

The invention includes methods for treating patients suffering from conditions characterized by a degradation of aggrecan or methods for preventing such conditions. These methods, according to the invention, entail administering to a patient needing such treatment, an effective amount of a composition comprising an aggrecanase inhibitor that inhibits the proteolytic activity of aggrecanase enzymes.

Still a further aspect of the invention is DNA sequences coding for expression of aggrecanase protein. Such sequences include the sequence of nucleotides in a 5' to 3' direction as illustrated in SEQ ID NO: 7. Further included in the present invention are DNA sequences which hybridize under stringent conditions with the DNA sequence of SEQ ID NO: 7 and encode a protein having the ability to cleave aggrecan. Preferred DNA sequences include those which hybridize under stringent conditions [see, Maniatis, T. et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory 387-389 (1982)]. It is generally preferred that such DNA sequences encode a polypeptide, which is at least about 80% homologous, and more

preferably at least about 90% homologous to the sequence as set forth in SEQ ID NO: 8. Finally, allelic or other variations of the sequence of SEQ ID NO: 7, whether such nucleotide changes result in changes in the peptide sequence or not, but where the peptide sequence still has aggrecanase activity, are also included in the present invention. The present invention also includes fragments of the DNA sequence as shown in SEQ ID NO: 7, which encodes a polypeptide that retains the activity of aggrecanase.

The DNA sequences of the present invention are useful, for example, as probes for the detection of mRNA encoding aggrecanase in a given cell population. Thus, the present invention includes methods of detecting or diagnosing genetic disorders involving the aggrecanase; or disorders involving cellular, organ, or tissue disorders in which aggrecanase is irregularly transcribed or expressed. The DNA sequences may also be useful for preparing vectors for gene therapy applications as described below.

A further aspect of the invention includes vectors comprising a DNA sequence as described above in operative association with an expression control sequence therefor. These vectors may be employed in a novel process for producing an aggrecanase protein of the invention in which a cell line transformed with a DNA sequence encoding an aggrecanase protein in operative association with an expression control sequence therefor is cultured in a suitable culture medium and an aggrecanase protein is recovered and purified therefrom. This process may employ a number of known cells, both prokaryotic and eukaryotic, as host cells for expression of the polypeptide. The vectors may be used in gene therapy applications. In such use, the vectors may be transfected into the cells of a patient *ex vivo*, and the cells may be reintroduced into a patient. Alternatively, the vectors may be introduced into a patient *in vivo* through targeted transfection.

Still a further aspect of the invention is aggrecanase proteins or polypeptides. Such polypeptides are characterized by having an amino acid sequence including the sequence illustrated in SEQ ID NO: 8, variants of the amino acid sequence of SEQ ID NO: 8, including naturally occurring allelic variants, and other variants in which the protein retains the ability to cleave aggrecan characteristic of aggrecanase molecules. Preferred polypeptides include a polypeptide, which is at least about 80% homologous, and more preferably at least about 90% homologous, to the amino acid sequence shown in SEQ ID NO: 8. Finally, allelic or other variations of the

sequences of SEQ ID NO: 8, whether such amino acid changes are induced by mutagenesis, chemical alteration, or by alteration of a DNA sequence used to produce the polypeptide, where the peptide sequence still has aggrecanase activity, are also included in the present invention. The present invention also includes fragments of the amino acid sequence of SED ID NO: 8, which retain the activity of aggrecanase protein.

The purified protein of the present invention may be used to generate antibodies, either monoclonal or polyclonal, to aggrecanase and/or other aggrecanase-related proteins, using methods that are known in the art of antibody production. Thus, the present invention also includes antibodies to aggrecanase or other related proteins. The antibodies may be useful for detection and/or purification of aggrecanase or related proteins, or for inhibiting or preventing the effects of aggrecanase. The aggrecanase of the invention or portions thereof may be utilized to prepare antibodies that specifically bind to aggrecanase.

DETAILED DESCRIPTION OF THE INVENTION

The human aggrecanase of the present invention comprises nucleotides #1 through #4284 as set forth in SEQ ID NO: 7. The human aggrecanase protein sequence is set forth in SEQ ID NO: 8 from amino acids #1 through #1427.

The aggrecanase protein of the present invention includes polypeptides comprising the amino acid sequence of SEQ ID NO: 8 having the ability to cleave aggrecan.

The aggrecanase proteins recovered from the culture medium are purified by isolating them from other proteinaceous materials from which they are co-produced and from other contaminants present. The isolated and purified proteins may be characterized by the ability to cleave aggrecan substrate. The aggrecanase proteins provided herein also include factors encoded by the sequences similar to those of SEQ ID NO: 7, but into which modifications or deletions are naturally provided (e.g. allelic variations in the nucleotide sequences which may result in amino acid changes in the polypeptide) or deliberately engineered. For example, synthetic polypeptides may wholly or partially duplicate continuous sequences of the amino acid residues of SEQ ID NO: 8. These sequences, by virtue of sharing primary, secondary, or tertiary structural and conformational characteristics with aggrecanase molecules may possess biological properties in common therewith.

It is known, for example, that numerous conservative amino acid substitutions are possible without significantly modifying the structure and conformation of a protein, thus maintaining the biological properties as well. For example, it is recognized that conservative amino acid substitutions may be made among amino acids with basic side chains, such as lysine (Lys or K), arginine (Arg or R) and histidine (His or H); amino acids with acidic side chains, such as aspartic acid (Asp or D) and glutamic acid (Glu or E); amino acids with uncharged polar side chains, such as asparagines (Asn or N), glutamine (Gln or Q), serine (ser or S), threonine (Thr or T), and tyrosine (Tyr or Y); and amino acids with nonpolar side chains, such as alanine (Ala or A), glycine (Gly or G), valine (Val or V), leucine (Leu or L), isoleucine (Ile or I), proline (Pro or P), phenylalanine (Phe or F), methionine (Met or M), tryptophan (Trp or W) and cysteine (Cys or C). Thus, these modifications and deletions of the native aggrecanase may be employed as biologically active substitutes for naturally occurring aggrecanase and in the development of inhibitors to other polypeptides in therapeutic processes. It can be readily determined whether a given variant of aggrecanase maintains the biological activity of aggrecanase by subjecting both aggrecanase and the variant of aggrecanase, as well as inhibitors thereof, to the assays described in the examples.

Other specific mutations of the sequences of aggrecanase proteins described herein involve modifications of glycosylation sites. These modifications may involve O-linked or N-linked glycosylation sites. For instance, the absence of glycosylation or only partial glycosylation results from amino acid substitution or deletion at asparagine-linked glycosylation recognition sites. The asparagine-linked glycosylation recognition sites comprise tripeptide sequences, which are specifically recognized by appropriate cellular glycosylation enzymes. These tripeptide sequences are either asparagine-X-threonine or asparagine-X-serine, where X is usually any amino acid. A variety of amino acid substitutions or deletions at one or both of the first or third amino acid positions of a glycosylation recognition site (and/or amino acid deletion at the second position) results in non-glycosylation at the modified tripeptide sequence. Additionally, bacterial expression of aggrecanase-related protein will also result in production of a non-glycosylated protein even if the glycosylation sites are left unmodified.

The present invention also encompasses novel DNA sequences, free of association with DNA sequences encoding other proteinaceous materials, and coding

for expression of aggrecanase proteins. These DNA sequences include SEQ ID NO: 7 in a 5' to 3' direction and those sequences which hybridize thereto under stringent hybridization washing conditions [for example, 0.1X SSC, 0.1% SDS at 65°C; see, Maniatis, T. et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory 387-389 (1982)] and encode a protein having aggrecanase proteolytic activity.

Similarly, DNA sequences which code for aggrecanase proteins coded for by the sequence of SEQ ID NO: 7, or aggrecanase protein which comprises the amino acid sequence of SEQ ID NO: 8, but which differs in codon sequence due to the degeneracies of the genetic code or allelic variations also encode the novel factors described herein. Variations in the DNA sequence of SEQ ID NO: 7, which are caused by point mutations or by induced modifications (including insertion, deletion, and substitution) to enhance the activity, half-life, or production of the polypeptides encoded are also encompassed in the invention.

Another aspect of the present invention provides a novel method for producing aggrecanase proteins. The method of the present invention involves culturing a suitable cell line, which has been transformed with a DNA sequence encoding an aggrecanase protein of the invention, under the control of known regulatory sequences. The transformed host cells are cultured and the aggrecanase proteins recovered and purified from the culture medium. The purified proteins are substantially free from other proteins with which they are co-produced as well as from other contaminants.

Suitable cells or cell lines may be mammalian cells, such as Chinese Hamster Ovary (CHO) cells. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening, product production, and purification are known in the art. See, e.g. Gething and Sambrook, Nature, 293: 620-625 (1981), or alternatively, Kaufman et al., Mol. Cell. Biol., 5(7): 1750-1759 (1985) or Howley et al., U.S. Patent 4,419,446. Another suitable mammalian cell line, which is described in the accompanying examples, is the monkey COS-1 cell line. The mammalian cell CV-1 may also be suitable.

Bacterial cells may also be suitable hosts. For example, the various strains of E. coli (e.g., HB101, MC1061) are well known as host cells in the field of biotechnology. Various strains of B. subtilis, Pseudomonas, other bacilli, and the like

may also be employed in this method. For expression of the protein in bacterial cells, DNA encoding the propeptide of aggrecanase is generally not necessary.

Many strains of yeast cells known to those skilled in the art may also be available as host cells for expression of the polypeptides of the present invention. Additionally, where desired, insect cells may be utilized as host cells in the method of the present invention. See, e.g. Miller et al., Genetic Engineering, 8: 277-298 (Plenum Press 1986) and references cited therein.

Another aspect of the present invention provides vectors for use in the method of expression of these novel aggrecanase polypeptides. Preferably, the vectors contain the full novel DNA sequences described above which encode the novel factors of the invention. Additionally, the vectors contain appropriate expression control sequences permitting expression of the aggrecanase protein sequences. Alternatively, vectors incorporating modified sequences as described above are also embodiments of the present invention. Additionally, the sequence of SEQ ID NO: 7 could be manipulated to express composite aggrecanase molecules. Thus, the present invention includes chimeric DNA molecules encoding an aggrecanase protein comprising a fragment from SEQ ID NO: 7 linked in correct reading frame to a DNA sequence encoding another aggrecanase polypeptide.

The vectors may be employed in the method of transforming cell lines and contain selected regulatory sequences in operative association with the DNA coding sequences of the invention which are capable of directing the replication and expression thereof in selected host cells. Regulatory sequences for such vectors are known to those skilled in the art and may be selected depending upon the host cells. Such selection is routine and does not form part of the present invention.

Various conditions such as osteoarthritis are known to be characterized by degradation of aggrecan. Therefore, an aggrecanase protein of the present invention, which cleaves aggrecan may be useful for the development of inhibitors of aggrecanase. The invention therefore provides compositions comprising an aggrecanase inhibitor. The inhibitors may be developed using the aggrecanase in screening assays involving a mixture of aggrecan substrate with the inhibitor followed by exposure to aggrecan. The compositions may be used in the treatment of osteoarthritis and other conditions exhibiting degradation of aggrecan. The invention further includes antibodies, which can be used to detect aggrecanase and also may be used to inhibit the proteolytic activity of aggrecanase.

The therapeutic methods of the invention include administering the aggrecanase inhibitor compositions topically, systemically, or locally as an implant or device. The dosage regimen will be determined by the attending physician considering various factors which modify the action of the aggrecanase protein including the site of pathology; the severity of disease; the patient's age, sex, and diet; the severity of any inflammation; time of administration; and other clinical factors. Generally, systemic or injectable administration will be initiated at a dose that is minimally effective, and the does will be increased over a preselected time course until a positive effect is observed. Subsequently, incremental increases in dosage will be made limiting such incremental increases to such levels that produce a corresponding increase in effect, while taking into account any adverse effects that may appear. The addition of other known factors to the final composition may also affect the dosage.

Progress can be monitored by periodic assessment of disease progression. The progress can be monitored, for example, by x-rays, MRI or other imaging modalities, synovial fluid analysis, and/or clinical examination.

The following examples illustrate practice of the present invention in isolating and characterizing human aggrecanase and other aggrecanase-related proteins, obtaining the human proteins and expressing the proteins via recombinant techniques.

EXAMPLES

EXAMPLE 1

Isolation of DNA

Potential novel aggrecanase family members were identified using a database screening approach. Aggrecanse-1 [Science 284: 1664-1666 (1999)] has at least six domains: signal, propeptide, catalytic domain, disintegrin, tsp, and c-terminal. The catalytic domain contains a zinc-binding signature region, TAAHELGHVKF (SEQ ID NO: 9), and a MET turn, which are responsible for protease activity. Substitutions within the zinc-binding region in the number of the positions still allow protease activity, but the histidine (H) and glutamic acid (E) residues must be present. The thrombospondin domain of Aggrecanase-1 is also a critical domain for substrate recognition and cleavage. These two domains determine our classification of a novel aggrecanase family member.

The protein sequence of the Aggrecanase-1 DNA sequence was used to query against the GenBank ESTs focusing on human ESTs using TBLASTN. The resulting sequences were the starting point in the effort to identify full-length sequences for potential family members. The nucleotide sequence of the aggrecanase of the present invention is comprised of five ESTs that contain homology over the catalytic domain and zinc-binding motif of Aggrecanase-1.

This human aggrecanase sequence was isolated from a dT-primed cDNA library constructed in the plasmid vector pED6-dpc2. The cDNA was made from human stomach RNA purchased from Clontech. The probe to isolate the aggrecanase of the present invention was generated from the sequence obtained from the database search. The sequence of the probe was as follows: 5'-GTGAGGTTGGCTGTGATATTTGGAGCAC-3' (SEQ ID NO: 10). The DNA probe was radioactively labeled ^{32}P and was used to screen the human stomach dT-primed cDNA library, under high stringency hybridization/washing conditions, to identify clones containing sequences of the human candidate #5.

Fifty thousand library transformants were plated at a density of approximately 5000 transformants per plate on 10 plates. Nitrocellulose replicas of the transformed colonies were hybridized to the ^{32}P labeled DNA probe in standard hybridization buffer (1X Blotto [25X Blotto = 5% nonfat dried milk, 0.02% azide in dH2O)] + 1% NP-40 + 6X SSC + 0.05% Pyrophosphate) under high stringency conditions (65°C for 2 hours). After 2 hours of hybridization, the radioactively labeled DNA probe containing hybridization solution was removed and the filters were washed under high stringency conditions (3X SSC, 0.05% Pyrophosphate for 5 minutes at RT; followed by 2.2X SSC, 0.05% Pyrophophate for 15 minutes at RT; followed by 2.2X SSC, 0.05% Pyrophosphate for 1-2 minutes at 65°C). The filters were wrapped in Saran wrap and exposed to X-ray film for overnight. The autoradiographs were developed and positively hybridizing transformants of various signal intensities were identified. These positive clones were picked, grown for 12 hours in selective medium, and plated at low density (approximately 100 colonies per plate). Nitrocellulose replicas of the colonies were hybridized to the ^{32}P labeled probe in standard hybridization buffer ((1X Blotto [25X Blotto = 5% nonfat dried milk, 0.02% azide in dH2O] + 1% NP-40 + 6X SSC + 0.05% Pyrophosphate) under high stringency conditions (65°C for 2 hours). After 2 hours of hybridization, the radioactively labeled DNA probe containing hybridization solution was removed and the filters were washed under high

stringency conditions (3X SSC, 0.05% Pyrophosphate for 5 minutes at RT; followed by 2.2X SSC, 0.05% Pyrophosphate for 15 minutes at RT; followed by 2.2X SSC, 0.05% Pyrophosphate for 1-2 minutes at 65°C). The filters were wrapped in Saran wrap and exposed to x-ray film for overnight. The autoradiographs were developed and positively hybridizing transformants were identified. Bacterial stocks of purified hybridization positive clones were made and plasmid DNA was isolated. The sequence of the cDNA insert was determined and is set forth in SEQ ID NO: 2 and 3. This sequence has been deposited in the American Type Culture Collection 10801 University Blvd. Manassas, VA 20110-2209 USA as PTA-2285. The cDNA insert contained the sequences of the DNA probe used in the hybridization.

The obtained human candidate #5 sequence aligns with several ESTs in the public database, along with a human cDNA, Hsa011374. Hsa011374 extends the aggrecanase sequence of the present invention about 2 kB at the 3' end. When two gaps are inserted into the Hsa011374 sequence, the aggrecanase sequence of the present invention can be lined up to create a sequence that is about 40% homologous to Aggrecanase–1. The aggrecanase of the present invention contains the zinc-binding region signature and a MET turn, however, is missing the signal and propeptide regions. The Hsa011374 extends our sequence to cover the disintegrin, tsp, and c-terminal spacer. It is with these criteria that candidate #5 is considered a novel aggrecanase family member.

This aggrecanase sequence of the invention can be used to design probes for further screening for full-length clones containing the isolated sequence. Based on the nucleotide sequences, numerous PCR primers were designed. The primers were used for both 3' and 5' Rapid Amplification of cDNA Ends (RACE) reactions and to amplify internal segments of the gene. All the amplified PCR products were cloned into vectors and sequenced. The computer program DNASTAR was used to align all the overlapping products and a consensus sequence was determined. Based on this new virtual DNA sequence, additional PCR primers were designed for the full-length cloning of the gene.

An OriGene Multi-Tissue RACE panel (HSCA-101) was screened to identify potential tissue sources for future experiments. The antisense primer 5'-CGCTACCTGAGCAGGCTCAGCAGCT-3' (SEQ ID NO: 11) was used with Clontech Advanatge GC2 polymerase reagents according to the manufacturer's recommendations. All amplifications were carried out in a Perkin-Elmer 9600

thermocycler. Cycling parameters were 94°C for 3 min, 5 cycles at 94°C for 30 sec, 65°C for 30 sec, 72°C for 5 min, 15 cycles at 94°C for 30 sec, 62°C for 30 sec, 72°C for 5 min, and 72°C for 6 min. First round reactions were diluted 10-fold with dH_2O, then 1 μl of the diluted first round reaction was used as a template for a second round of amplification with the nested primer 5'-CCCGAAGCAGTTCTGCCCCGATGTTG-3' (SEQ ID NO: 12) utilizing the identical parameters as described for the first round. 10 μl of the second round reaction was fractionated on 1% agarose gel and then transferred to nitrocellulose for Southern analysis. The nitrocellulose membrane was prehybridized in Clontech ExpressHyb for 30 minutes at 37°C according to the manufacture recommendations. The membrane was then incubated with $1x10^6$ CPM of the γ-ATP end-labeled oligo 5'-ACCCGAGTTGTCTTCAGGCTTTGGA-3' (SEQ ID NO: 13) at 37°C for 1 hour. Unbound probe was removed by two washes at room temperature with 2x SSC/0.05% SDS followed by two additional washes at room temperature with 0.1x SSC/0.1% SDS. Autoradiography suggested human candidate #5 was present in tissues including: testis, stomach, liver, heart, and colon.

Liver Marathon-Ready cDNA (Clontech) was used as a template in PCR cloning reactions. The antisense primer 5'-CTCCACGCTTCATGATGAAGCTCTCG-3' (SEQ ID NO: 14) was used in a first round 5' RACE reaction and the sense primer 5'-GCGGCGCCTCCTTCTACCACT-3' (SEQ ID NO: 15) was used in the first round 3' RACE reaction. Clontech Advantage GC2 polymerase reagents were used according to the manufacture recommendations. All amplifications were carried out in a Perkin-Elmer 9600 thermocycler. Cycling parameters were 94°C for 30 secs, 5 cycles of 94°C for 5 sec, 72°C for 4 min, 5 cycles of 94°C for 5 sec, 70°C for 4 min, 30 cycles at 94°C for 5 sec, and 68°C for 4 min. The first round reactions were diluted 10 fold in TE and 5 μl was used as template for a second round of PCR. The antisense primer 5'-TCCGTGTCGTCCTCAGGGTTGATGG-3' (SEQ ID NO: 16) or 5'-CCCTCAGGCTCTGTCAGAATGACCA-3' (SEQ ID NO: 17) was used for second round 5' RACE and the sense primer 5'-AGGGGCCTGGCTCCGTAGATG-3' (SEQ ID NO: 18) or

5'-CTGGGAGCCGGCGGGAGGTCTGC-3' (SEQ ID NO: 19) was used for second round 3' RACE utilizing the identical parameters as described for the first round. Aliquots of each reaction were fractionated on a 1% agarose gel and the oligos 5'-CCACAGGCCGTGTCTTCTTACTTGA-3' (SEQ ID NO: 20) and 5'-CCATGGGCCCGGGCACAATACAGG-3' (SEQ ID NO: 21) were end labeled and used as probes for Southern analysis of the 5' and 3' RACE products, respectively. Conditions for Southern analysis were as described above. Duplicate agarose gels were run and the PCR products that corresponded with positive signals on the autorads were cut out of the agarose gel and the DNA was recovered from the gel matrix via BioRad's Prep-A-Gene DNA Purification System. The recovered DNA was ligated into either Clontech's Advantage PCR cloning kit or Stratagene's PCR-Script Amp Cloning Kit according to the manufacturer's instructions. Vectors were transformed into Life Technologies ElectroMax DH10B cells according to the manufacturer's recommendations.

The primer pair 5'-CAACATCGGGGCAGAACTGCTTCGGG-3' (SEQ ID NO: 22) and 3'-CCATGGGCCCGGGCACAATACAGG-5' (SEQ ID NO: 23) were used in conjunction with Clontech Liver Marathon-Ready cDNA to amplify an internal 2622 bp fragment of EST5. PCR cycling conditions and reagents were identical to conditions used for the RACE reactions. The 2622 bp fragment was cloned into the PCR-Script vector as described above.

Assembly of all the cloned fragments in DNASTAR produced a single ORF of 4284 bp. The full-length cloning of the gene was then accomplished by amplifying three overlapping DNA fragments, digesting the fragments with specific restriction enzymes followed by ligation and transformation into DH10B cells. Stratagene's Pfu Turbo Hotstart DNA polymerase was used to amplify each fragment from Clontech Liver Marathon-Ready cDNA. In addition to following conditions recommended by the manufacturer's DMSO was included at a final concentration of 5% in each PCR reaction. Cycling parameters were 94°C for 30 sec, 5 cycles of 94°C for 5 sec, 72°C for 4 min, 5 cycles at 94°C for 5 sec, 70°C for 4 min, 30 cycles at 94°C for 5 sec, and 68°C for 4 min. The following are primer pairs used to amplify each fragment:

	PCR product (base pairs)	
	undigested	digested
Fragment 1	1833 bp	717 bp
5'-TAAATCGAATTCCCACCATGCACCAGCGTCACCCCTGGGCA-3' (SEQ ID NO: 24)		
3'-CCACGACATAGCGCCCTCCGATCCT-5' (SEQ ID NO: 25)		
Fragment 2	2622 bp	2211 bp
5'-CAACATCGGGGCAGAACTGCTTCGGG-3' (SEQ ID NO: 26)		
3'-CCATGGGCCCGGGCACAATACAGG-5' (SEQ ID NO: 27)		
Fragment 3	1770 bp	1754 bp
5'-AGGGGCCTGGCTCCGTAGATG-3' (SEQ ID NO: 28)		
3'-ATAGTTTAGCGGCCGCTCAGGTTCCTTCCTTTCCCTTCCAG-5' (SEQ ID NO: 28)		

	EcoRI	AscI	BamHI	NotI
	↓	↓	↓	↓
fragment 1	----------			
fragment 2		----------		
fragment 3			----------	

PCR products were digested with the indicated enzymes and then fractionated on 1% agarose gel. DNA bands corresponding to the indicated digested sizes were recovered from the gel as described above. The ligation reaction included equal molar ratios of the three digested DNA fragments and the vector pHTOP pre-digested EcoRI-NotI. The full-length gene construction was confirmed by DNA sequencing and is set forth in SEQ ID NO: 7, and the amino acid sequence is set forth in SEQ ID NO: 8.

EXAMPLE 2

Expression of Aggrecanase

In order to produce murine, human, or other mammalian aggrecanase-related proteins, the DNA encoding it is transferred into an appropriate expression vector and introduced into mammalian cells or other preferred eukaryotic or prokaryotic hosts including insect host cell culture systems by conventional genetic engineering

techniques. The expression system for biologically active recombinant human aggrecanase is contemplated to be stably transformed mammalian, insect, yeast, or bacterial cells.

One skilled in the art can construct mammalian expression vectors by employing the sequence of SEQ ID NO: 7 and known vectors, such as pCD [Okayama et al., Mol. Cell Biol., 2:162-170 (1982)], pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)], and pMT2 CXM.

The mammalian expression vector pMT2 CXM is a derivative of p91023(b) (Wong et al., Science 228: 810-815, 1985) differing from the latter in that it contains the ampicillin resistance gene in place of the tetracycline resistance gene and further contains a XhoI site for insertion of cDNA clones. The functional elements of pMT2 CXM have been described (Kaufman, R.J., Proc. Natl. Acad. Sci. USA 82: 689-693 (1985)) and include the adenovirus VA genes, the SV40 origin of replication including the 72 bp enhancer, the adenovirus major late promoter including a 5' splice site and the majority of the adenovirus tripartite leader sequence present on adenovirus late mRNAs, a 3' splice acceptor site, a DHFR insert, the SV40 early polyadenylation site (SV40), and pBR322 sequences needed for propagation in E coli.

Plasmid pMT2 CXM is obtained by EcoRI digestion of pMT2-VWF, which has been deposited with the American Type Culture Collection (ATCC), Rockville, MD (USA) under accession number ATCC 67122. EcoRI digestion excises the cDNA insert present in pMT2-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods. pMT2 CXM is then constructed using loopout/in mutagenesis [Morinaga et al., Biotechnology 84: 636 (1984). This removes bases 1075 to 1145 relative to the Hind III site near the SV40 origin of replication and enhancer sequences of pMT2. In addition it inserts the following sequence:
5'-CATGGGCAGCTCGAG-3' (SEQ ID NO: 30) at nucleotide 1145. This sequence contains the recognition site for the restriction endonuclease XhoI. A derivative of pMT2 CXM , termed pMT23, contains recognition sites for the restriction endonucleases PstI, Eco RI, SalI, and XhoI. Plasmid pMT2 CXM and pMT23 DNA may be prepared by conventional methods.

pEMC2β1 derived from pMT21 may also be suitable in the practice of the invention. pMT21 is derived from pMT2 which is derived from pMT2-VWF. As

described above, EcoRI digestion excises the cDNA insert present in pMT-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods.

pMT21 is derived from pMT2 through the following two modifications. First, 76 bp of the 5' untranslated region of the DHFR cDNA including a stretch of 19 G residues from G/C tailing for cDNA cloning is deleted. In this process, a XhoI site is inserted to obtain the following sequence immediately upstream from DHFR:

5'-CTGCAGGCGAGCCTGAATTCCTCGAGCCATCATG-3' (SEQ ID NO: 31)

PstI EcoRI XhoI

Second, a unique ClaI site is introduced by digestion with EcoRV and XbaI, treatment with Klenow fragment of DNA polymerase I, and ligation to a ClaI linker (CATCGATG). This deletes a 250 bp segment from the adenovirus associated RNA (VAI) region but does not interfere with VAI RNA gene expression or function. pMT21 is digested with EcoRI and XhoI, and is used to derive the vector pEMC2B1.

A portion of the EMCV leader is obtained from pMT2-ECAT1 [Jung, S.K. et al., J. Virol 63: 1651-1660 (1989)] by digestion with EcoRI and PstI, resulting in a 2752 bp fragment. This fragment is digested with TaqI yielding an EcoRI-TaqI fragment of 508 bp which is purified by electrophoresis on a low melting agarose gel. A 68 bp adapter and its complementary strand are synthesized with a 5' TaqI protruding end and a 3' XhoI protruding end which has the following sequence: 5'-

CGAGGTTAAAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTC

TaqI

CTTTGAAAAACACGATTGC-3' (SEQ ID NO: 32)

XhoI

This sequence matches the EMC virus leader sequence from nucleotide 763 to 827. It also changes the ATG at position 10 within the EMC virus leader to an ATT and is followed by a XhoI site. A three-way ligation of the pMT21 EcoRIXhoI fragment, the EMC virus EcoRI-TaqI fragment, and the 68 bp oligonucleotide adapter TaqI-XhoI adapter resulting in the vector pEMC2β1.

This vector contains the SV40 origin of replication and enhancer, the adenovirus major late promoter, a cDNA copy of the majority of the adenovirus tripartite leader sequence, a small hybrid intervening sequence, an SV40 polyadenylation signal and the adenovirus VA I gene. DHRF and β-lactamase

markers, and an EMC sequence in appropriate relationships to direct the high level expression of the desired cDNA in mammalian cells.

The construction of vectors may involve modification of the aggrecanase-related DNA sequences. For instance, aggrecanase cDNA can be modified by removing the non-coding nucleotides on the 5' and 3' ends of the coding region. The deleted non-coding nucleotides may or may not be replaced by other sequences known to be beneficial for expression. These vectors are transformed into appropriate host cells for expression of aggrecanase-related proteins. Additionally, the sequence of SEQ ID NO: 7 can be manipulated to express a mature aggrecanase-related protein by deleting aggrecanase encoding propeptide sequences and replacing them with sequences encoding the complete propeptides of other aggrecanase proteins.

One skilled in the art can manipulate the sequence of SEQ ID NO: 7 by eliminating or replacing the mammalian regulatory sequences flanking the coding sequence with bacterial sequences to create bacterial vectors for intracellular or extracellular expression by bacterial cells. For example, the coding sequences could be further manipulated (e.g. ligated to other known linkers or modified by deleting non-coding sequences therefrom or altering nucleotides therein by other known techniques). The modified aggrecanase-related coding sequence could then be inserted into a known bacterial vector using procedures such as described in Taniguchi, T. et al., Proc. Natl Acad. Sci. USA, 77:5230-5233 (1980). This exemplary bacterial vector could then be transformed into bacterial host cells and an aggrecanase-related protein expressed thereby. For a strategy for producing extracellular expression of aggrecanase-related proteins in bacterial cells, see, e.g. European patent application EPA 177,343.

A method for producing high levels of an aggrecanase-related protein of the invention in mammalian, bacterial, yeast, or insect host cell systems may involve the construction of cells containing multiple copies of the heterologous aggrecanase-related gene. The heterologous gene is linked to an amplifiable marker, e.g. the DiHydroFolate Reductase (DHFR) gene for which cells containing increased gene copies can be selected for propagation in increasing concentrations of MethoTreXate (MTX) according to the procedures of Kaufman and Sharp, J. Mol. Biol., 159:601-629 (1982). This approach can be employed with a number of different cell types.

For example, a plasmid containing a DNA sequence for an aggrecanase-related protein of the invention in operative association with other plasmid sequences

enabling expression thereof and the DHFR expression plasmid pAdA26SV(A)3 [Kaufman and Sharp, Mol. Cell. Biol., 2: 1304 (1982)] can be co-introduced into DHRF-deficient CHO cells, DUKX-BII, by various methods including calcium phosphate coprecipitation and transfection, electroporation or protoplast fusion DHRF expressing transformants are selected for growth in alpha media with dialyzed fetal calf serum, and subsequently selected for amplification by growth in increasing concentrations of MTX (e.g. sequential steps in 0.02, 0.2, 1.0 and 5 uM MTX) as described in Kaufman et al., Mol Cell Biol., 5: 1750 (1983). Transformants are cloned, and biologically active aggrecanase expression is monitored by the assays described above. Aggrecanase protein expression should increase with increasing levels of MTX resistance. Aggrecanase polypeptides are characterized using standard techniques known in the art such as pulse labeling with [35S] methionine or cysteine and polyacrylamide gel electrophoresis. Similar procedures can be followed to produce other related aggrecanase-related proteins.

As one example, the aggrecanase gene of the present invention is cloned into the expression vector pED6 [Kaufman et al., Nucleic Acid Res. 19: 4485-4490 (1991)]. COS and CHO DUKX B11 cells are transiently transfected with the aggrecanase sequence of the invention (+/- co-transfection of PACE on a separate pED6 plasmid) by lipofection (LF2000, Invitrogen). Duplicate transfections are performed for each gene of interest: (a) one for harvesting conditioned media for activity assay and (b) one for 35S-methionine/cysteine metabolic labeling.

On day one media is changed to DME (COS) or alpha (CHO) media-1% heat-inactivated fetal calf serum +/- 100 μg/ml heparin on wells (a) to be harvested for activity assay. After 48h (day 4), conditioned media is harvested for activity assay.

On day 3, the duplicate wells (b) are changed to MEM (methionine-free/cysteine free) media +1% heat-inactivated fetal calf serum +100 μg/ml heparin + 100 μCi/ml 35S-methionine/cysteine (Redivue Pro mix, Amersham). Following 6h incubation at 37°C, conditioned media is harvested and run on SDS-PAGE gels under reducing conditions. Proteins are visualized by autoradiography.

EXAMPLE 3

Biological Activity of Expressed Aggrecanase

To measure the biological activity of the expressed aggrecanase-related proteins obtained in Example 2, above, the proteins are recovered from the cell culture and purified by isolating the aggrecanase-related proteins from other proteinaceous materials with which they are co-produced as well as from other contaminants. The purified protein may be assayed in accordance with assays described above. Purification is carried out using standard techniques known to those skilled in the art.

Protein analysis is conducted using standard techniques such as SDS-PAGE acrylamide [Laemmli, Nature 277: 680 (1970)] stained with silver [Oakley, et al. Anal. Biochem. 105: 361 (1980)] and by immunoblot [Towbin et al. Proc. Natl. Acad. Sci. USA 76: 4350 (1979)].

The foregoing descriptions detail presently preferred embodiments of the present invention. Numerous modifications and variations in practice thereof are expected to occur to those skilled in the art upon consideration of these descriptions. Those modifications and variations are believed to be encompassed within the claims appended hereto.